Effect of Doublon-Holon Binding on Mott transition—Variational Monte Carlo Study of Two-Dimensional Bose Hubbard Models

Hisatoshi Yokoyama[*], Tomoaki Miyagawa and Masao Ogata[1]

Department of Physics, Tohoku University, Sendai 980-8578
[1]*Department of Physics, University of Tokyo, Bunkyo-ku, Tokyo 113-0033*

To understand the mechanism of Mott transitions in case of no magnetic influence, superfluid-insulator (Mott) transitions in the $S = 0$ Bose Hubbard model at unit filling are studied on the square and triangular lattices, using a variational Monte Carlo method. In trial many-body wave functions, we introduce various types of attractive correlation factors between a doubly-occupied site (doublon, D) and an empty site (holon, H), which play a central role for Mott transitions, in addition to the onsite repulsive (Gutzwiller) factor. By optimizing distance-dependent parameters, we study various properties of this type of wave functions. With a hint from the Mott transition arising in a completely D-H bound state, we propose an improved picture of Mott transitions, by introducing two characteristic length scales, the D-H binding length ξ_{dh} and the minimum D-D exclusion length ξ_{dd}. Generally, a Mott transition occurs when ξ_{dh} becomes comparable to ξ_{dd}. In the conductive (superfluid) state, domains of D-H pairs overlap with each other ($\xi_{dh} > \xi_{dd}$); thereby D and H can propagate independently as density carriers by successively exchanging the partners. In contrast, intersite repulsive Jastrow (D-D and H-H) factors have little importance for the Mott transition.

KEYWORDS: Mott transition, Bose Hubbard model, two dimensions, superfluid, insulator, variational Monte Carlo Method, doublon-holon binding

1. Introduction

After early studies of superfluid-insulator or Mott transitions in interacting Bose systems,[1] an experimental examples have been realized using ultracold dilute gases of bosonic atoms in various optical lattices.[2–5] The essence of these systems is captured by a spinless Bose Hubbard model with a harmonic confinement potential.[6,7] Aside from this one-body potential, this basic model is important on the theoretical ground that a Mott transition can be studied without cares of magnetic influence, unlike typical fermionic cases. Most researchers are certain that, for the Fermi Hubbard models, metal-insulator transitions take place at infinitesimal correlation strength on the hypercubic lattice in any dimension.[8] Such transitions at $U/t = 0$ (U: onsite-interaction strength, t: hopping integral) require elements other than competition between itineracy and bare interaction of particles, like magnetic correlation. In contrast, spinless Bose Hubbard models bring about Mott transitions at moderate finite values of U/t.[9–12]

To date, Mott transitions in the $S = 0$ Bose Hubbard model have been studied with various methods. For the square lattice of our interest, properties of T_c, superfluid density, etc. were studied, applying a quantum Monte Carlo (QMC) method earlier to small systems[13] and later to larger systems;[14,15] a ground-state phase diagram in the plane of chemical potential and interaction strength was constructed using a strong-coupling expansion.[16] These studies estimated the critical interaction strength of Mott transitions at $U_c/t = 16.4$-16.7 for the particle density of unit filling, $n = 1$ ($n = N/N_s$ with N:

particle number, N_s: site number) at $T = 0$. Thus, the existence of a Mott transition has been embodied, but the mechanism of the transition is still not clear.

Variational Monte Carlo (VMC) approaches[17–19] are very useful to analyze the physics of Mott transitions, because one can directly and exactly treat wave functions for any values of U/t. As variational approaches to the Bose Hubbard model, wave functions with only onsite correlation factors, which correspond to the celebrated Gutzwiller wave function[20] (GWF, Ψ_G) for fermions, were studied first.[21,22] In contrast to for fermions, GWF for bosons is solved analytically without additional mean-field-type approximations[23] in arbitrary dimensions, and yield a Brinkman-Rice-type (BR) superfluid-insulator transition[24] at finite U ($\equiv U_{BR}$, see Fig. 2). In the insulating side of U_{BR}/t, however, all the lattice sites are occupied with exactly one particle and the hopping completely ceases, namely, $\Psi_G \rightarrow \prod_{j=1}^{N} b_j^\dagger |0\rangle$ and the total energy vanishes ($E = 0$). This result, which apparently contradicts experiments and reliable theories like the strong-coupling expansion, is caused by an over-simplified setup of the wave function, in which the effect of density fluctuation should be included. To remedy this drawback, it is crucial to add appropriate intersite correlation factors to GWF. In this line, recent VMC studies[25] emphasized the importance of an ordinary type of long-range Jastrow factor for the transition.

In this paper, we first show that an attractive correlation factor between a doubly-occupied site (doublon: D) and an empty site (holon: H) plays a leading role to induce the Mott transition for the present bosonic model. In fact, D-H near-neighbor correlations have long been studied for Fermi Hubbard models,[26–29] and the present

[*]E-mail address: yoko@cmpt.phys.tohoku.ac.jp

authors have shown that near-neighbor D-H binding factors and their analogs are capable of inducing Mott transitions for attractive Hubbard models,[30] a repulsive Hubbard model on an extended square lattice[31, 32] and on an anisotropic triangular lattice.[33] In this mechanism, in the insulating regime, a doublon and a holon as density carriers are confined in the range of near-neighbor sites; namely, density fluctuation is localized. In this work, we extend the D-H binding correlation factor to various long-range types in order to corroborate the above conclusion, and propose a renewed picture of conduction, which can explain a Mott transition arising in a completely D-H bound state. Namely, the Mott transition occurs, when a D-H binding length $\xi_{\rm dh}$ regulated by the D-H attractive correlation becomes comparable to the minimum D-D (H-H) distance $\xi_{\rm dd}$. This corresponds to the condition that a D-H pair comes to stop exchanging the partner with nearby D-H pairs. Furthermore, we introduce a repulsive Jastrow factor to check that it is of no importance for the Mott transition. These properties for bosons are fundamentally common to fermions, unless magnetic correlation is explicitly introduced.[34]

This paper is organized as follows: In §2, the model and the trial wave functions used in this paper are introduced. In §3, the properties of the Bose Hubbard model as to the Mott transition are studied with a short-range D-H binding factor. In §4, we consider the effect of long-range D-H attractive correlation factors, and of D-D repulsive factors. In §5, we discuss a renewed picture of the Mott transition by introducing two characteristic length scales. In §6, a concise summary is given. In Appendix, we briefly note the setup and condition of the VMC calculations implemented in this paper.

Parts of the results in this paper have been published before.[35, 36]

2. Formalism

After defining the model in §2.1, we introduce trial wave functions with various types of D-H attractive correlation factors and a long-range D-D (and H-H) repulsive factor in §2.2.

2.1 Bose Hubbard Model

We consider the $S = 0$ Bose Hubbard model on the square (SQL) and triangular (TAL) lattices with only homogeneous nearest-neighbor (NN) hopping:

$$H = -t \sum_{\langle ij \rangle} (b_i^\dagger b_j + b_j^\dagger b_i) + \frac{U}{2} \sum_j n_j(n_j - 1), \quad (1)$$

where b_j $(n_j = b_j^\dagger b_j)$ denotes an annihilation (number) operator of a boson on the site j. We assume $t, U \geq 0$, and use t as the energy unit. In this paper, we disregard a harmonic confinement potential, not only because we allow for the comparison with electron systems in solids, but because flat confinement potentials will be expected in forthcoming experiments of cold atoms.[37] We use systems of $L \times L$ $(= N_{\rm s})$ sites with the periodic boundary conditions in both x and y directions. The kinetic part,

H_t, in eq. (1) is diagonalized as,

$$H_t = \sum_{\bf k} \varepsilon_{\bf k} b_{\bf k}^\dagger b_{\bf k}, \quad (2)$$

with

$$\varepsilon_{\bf k} = \begin{cases} -2t(\cos k_x + \cos k_y), & {\rm SQL} \\ -2t\left[\cos k_x + \cos k_y + \cos(k_x + k_y)\right], & {\rm TAL} \end{cases} \quad (3)$$

using a Fourier transformation,

$$b_{\bf k} = \frac{1}{\sqrt{N_{\rm s}}} \sum_j e^{-i{\bf k}\cdot{\bf r}_j} b_j. \quad (4)$$

For $U/t = 0$, since all the particles condense into the lowest-energy level ${\bf k} = {\bf 0} = (0, 0)$, the ground state is

$$\Phi_0 = \frac{1}{\sqrt{N!}} \, b_{{\bf k}={\bf 0}}^{\dagger N} \, |0\rangle, \quad (5)$$

with the eigenenergy per site being $E = -4t$ $(-6t)$ for SQL (TAL). Using eq.(4), the real-space representation of Φ_0 is written as,

$$\Phi_0 \propto \sum_{\{{\bf r}\}} b_1^\dagger b_2^\dagger \cdots b_N^\dagger |0\rangle, \quad (6)$$

where $\{{\bf r}\}$ indicates the sum of all the particle configurations $\{{\bf r}_1, \cdots, {\bf r}_N\}$, permitting duplicate counting in combinations, and we abbreviate $b_{{\bf r}_i}^\dagger$ to $b_i^\dagger$. Note that, in contrast to fermionic ground states, the coefficient (a permanent of $N \times N$) of every configuration in the sum of eq. (6) becomes an identical constant, because all the elements in the permanent becomes unity. Accordingly, VMC calculations are greatly simplified.

As mentioned in §1, the ground-state phase diagram and some relevant properties of this model including the Mott critical value, $U_{\rm c}/t$, have been studied by reliable methods like QMC. Here, we shed light on the properties and mechanism as to the Mott transition, taking advantage of the VMC method.

2.2 Variational wave functions

We adopt a many-body variational approach to tackle the Mott physics in eq. (1). The simplest trial function is a Bose analog of the Gutzwiller wave function[20] (GWF), $\Psi_{\rm G} = g^D \Phi_0$, where g is a variational parameter controlling the number of multiply occupied sites (multiplon: M) and D is onsite correlation operator:

$$D = \frac{1}{2} \sum_j n_j(n_j - 1). \quad (7)$$

The correct energy expectation value by $\Psi_{\rm G}$ is given by an analog of the Gutzwiller approximation formula,[21, 22] and a Brinkman-Rice-type transition occurs at $U_{\rm BR}/t = (1 + \sqrt{2})^2 z$, where z is the number of the NN sites: $U_{\rm BR}/t = 23.31$ for SQL, and 34.97 for TAL. However, the description of the insulating state by $\Psi_{\rm G}$ is incorrect in that every site is occupied by a single particle and density fluctuation is completely suppressed, $\langle H_t \rangle = \langle H_U \rangle = 0$ $(H_U = UD)$, in the same way as the Brinkman-Rice transition in the fermionic cases.[24]

At unit filling, a multiplon and a holon are regarded as

positive and negative particle-density "carriers" respectively in the background of singly-occupied sites of the neutral or average particle density. Thus, conductivity depends on whether the motion of such density carriers is free or bound; to describe the Mott transition with higher fidelity, it is crucial to add an inter-carrier correlation factor, especially, a doublon(multiplon)-holon (D-H) factor, $\mathcal{P}_Q$:[29)]

$$\Psi_{\mathrm{DH}} = \mathcal{P}_Q \Psi_{\mathrm{G}}. \qquad (8)$$

For bosons, multiple site occupation of p (≥ 0) particles is allowed; for $U/t = 0$, the distribution of the number of sites occupied by p particles, $P(p)$, should be Poissonian. However, for large values of U/t like in a Mott critical region of our interest, $P(p)$ with $p \geq 3$ almost completely vanishes, as will be discussed later. Then, a multiplon becomes identical to a doublon, and the particle-hole symmetry is restored at $n = 1$. In this context, we treat multiplons and holons symmetrically in $\mathcal{P}_Q$, and often regard M as D for $U \gtrsim U_c/2$.

In our previous studies for fermions,[29–33)] we have used only near-neighbor D-H binding factors [(1) below], which are sufficient to describe Mott transitions. In this work, we introduce long-range types of $\mathcal{P}_Q$, whose necessity has long been recognized in an exact-diagonalization study in one dimension,[29)] to study the effect of D-H factor more in detail. We itemize $\mathcal{P}_Q$ used in this paper below.

(1) We extend the above near-neighbor D-H binding factors for fermions[38)] to the Bose Hubbard model. Following the previous papers, we call this short-range D-H binding wave function QWF or $\Psi_{\mathrm{DH}}^{\mu(\mu')}$. For SQL, we consider up to the second (diagonal)-neighbor correlation:

$$\mathcal{P}_Q = \mathcal{P}_Q(\mu, \mu') = (1 - \mu)^{\hat{Q}} (1 - \mu')^{\hat{Q}'} \qquad (9)$$

where primes indicate diagonal neighbors and,

$$\hat{Q}^{(')} = \sum_i \left[\tilde{d}_i \prod_{\tau^{(')}} (1 - h_{i+\tau^{(')}}) + h_i \prod_{\tau^{(')}} (1 - \tilde{d}_{i+\tau^{(')}}) \right]. \qquad (10)$$

Here, $\tilde{d}_i$ and h_i are projection operators of multiplon (M) and holon (H) on the site i, respectively:

$$\tilde{d}_i |i\rangle = \begin{cases} 1\,|i\rangle, & |i\rangle: \mathrm{M} \\ 0\,|i\rangle, & |i\rangle: \text{otherwise} \end{cases}, \qquad (11)$$

$$h_i |i\rangle = \begin{cases} 1\,|i\rangle, & |i\rangle: \mathrm{H} \\ 0\,|i\rangle, & |i\rangle: \text{otherwise} \end{cases}, \qquad (12)$$

$\tau^{(')}$ runs all the NN (diagonal-neighbor) sites of the site i, and $\mu^{(')}$ is a variational parameter controlling the binding of M and H between NN (diagonal-neighbor) sites. For TAL, we take account only of the NN M-H correlation:

$$\mathcal{P}_Q(\mu) = (1 - \mu)^{\hat{Q}}. \qquad (13)$$

In short, $\hat{Q}$ counts the number of M without NN H plus that of H without NN M. The range of μ is limited to $0 \leq \mu \leq 1$, whereas a subsidiary parameter μ' sometimes becomes negative.[39)] The density of M (H) isolated from

H (M) is reduced by μ; for $\mu = (\mu' =) 0$, Ψ_{DH}^μ is reduced to Ψ_{G}, and M can move freely from H. In the other limit, $\mu = 1$, M (H) cannot appear unless it is accompanied by at least one H (M) in the adjacent sites. Such a complete D-H bound state is not necessarily insulating, as we will see later. In Fig. 1, we show the weight of $\mathcal{P}_Q(\mu, 0)$ as a function of nearest M(D)-to-H distance r for an intermediate value ($\mu = 0.7$).
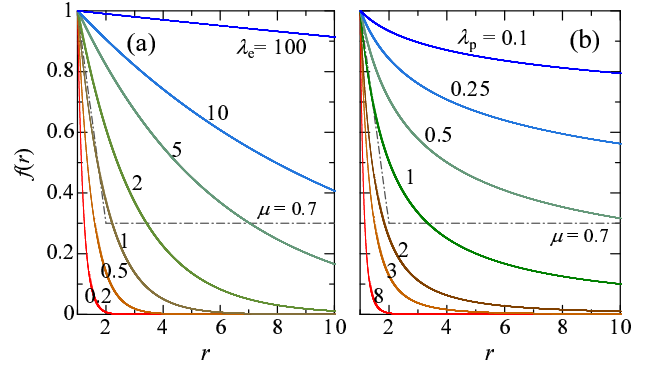

Fig. 1. (Color online) Weight of long-range doublon(D)-holon(H) correlation factor in (2) as a function of distance between D and the nearest H for several values of the parameter λ, eq.(15); (a) exponentially decaying type in $\Psi_{\mathrm{DH}}^{\mathrm{exp}}$, and (b) power-law decaying type in $\Psi_{\mathrm{DH}}^{\mathrm{pwr}}$. For comparison, $f(r)$ of the nearest-neighbor factor $\mathcal{P}_Q(0.7, 0)$ [eq.(9)] in Ψ_{DH}^μ is plotted with dash-dotted lines.

(2) We introduce long-range D-H binding factors of a simple form, which is formally written as,

$$\mathcal{P}_Q(\lambda) = \prod_j \left\{ \left[1 - (1 - f(r_j)) \tilde{d}_j \right] \left[1 - (1 - f(r_j)) h_j \right] \right\}, \qquad (14)$$

where λ is a variational parameter included in $f(r)$, which controls the size of an M-H pair. In $\mathcal{P}_Q(\lambda)$, we consider M-H correlation only between each M and its nearest H, and vice versa; r (≥ 1) denotes the distance between such M and H in unit of lattice constant, and is measured by the stepwise or "Manhattan" metric. This fashion suits the spirit of strong-coupling expansion, and is not identical with ordinary Jastrow factors, in which all the pairs are taken into account. For $f(r)$, on the analogy of a fermionic case,[29, 34)] we assume two primitive decaying forms, imposing $f(1) = 1$:

$$f(r) = \begin{cases} \exp\left(-\dfrac{r-1}{\lambda_{\mathrm{e}}}\right), & \text{(a) exponential} \\[2mm] \dfrac{1}{r^{\lambda_{\mathrm{p}}}}, & \text{(b) power} \end{cases} \qquad (15)$$

whose behavior is sketched in Fig.1. We write $\Psi_{\mathrm{DH}}^{\mathrm{exp}}$ and $\Psi_{\mathrm{DH}}^{\mathrm{pwr}}$ for the wave functions using the factors of eqs. (15a) and (15b), respectively. Compared with the exponential form (a), the power-law form (b) naturally has a tail for large r for intermediate to large λ_{e} and $1/\lambda_{\mathrm{p}}$. When λ_{e} or $1/\lambda_{\mathrm{p}} \to \infty$, $\mathcal{P}_Q(\lambda)$ becomes unity, namely, Ψ_{DH} is reduced to Ψ_{G}. Meanwhile, in the limit of λ_{e} or $1/\lambda_{\mathrm{p}} \to 0$, $\mathcal{P}_Q(\lambda)$ is reduced to $\mathcal{P}_Q(\mu = 1, \mu' = 0)$, indicating the complete D-H binding within the NN sites. Since the probability density of M-H pairs with distance r is con-

trolled mainly by g for small U/t and by λ for intermediate and large U/t, the mean M-H distance relates to λ_e or λ_p for $U \gtrsim U_c$.

(3) Instead of such an a priori form of $f(r)$ as eq. (15), we optimize $f(r)$ in eq. (14) for every r as variational parameters, using recently-developed optimization techniques.[19] In this case $[\mathcal{P}_Q(f)]$, the number of variational parameters becomes equivalent to the linear dimension of the system L: g and $f(2)$-$f(L)$. We represent this optimized-$f(r)$ wave function as $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$. As we will see later, parameters $f(r)$ with large r ($\gtrsim 7$), are actually unnecessary for any value of U/t, because such long-distance D-H pairs rarely appear. Although $\mathcal{P}_Q(f)$ is naturally better than both $\mathcal{P}_Q(\mu, \mu')$ and $\mathcal{P}_Q(\lambda)$, we can learn much from the comparison with various types of $\mathcal{P}_Q$.

Having considered an attractive part of the long-range correlation factor, we should check the effect on a Mott transition of a repulsive part, which is known to be important for off half filling.[29] To this end, we consider the form, $\Psi_J = \mathcal{P}_J \Psi_{\mathrm{DH}}$, with a common repulsive Jastrow factor,

$$\mathcal{P}_J(a, \kappa) = \prod_{j,r} \Big\{ \Big[1 - (1 - \eta(r))\, \tilde{d}_j \tilde{d}_{j+r} \Big]$$
$$\times \Big[1 - (1 - \eta(r))\, h_j h_{j+r} \Big] \Big\}, \quad (16)$$

in which we take account of the correlation between all the M-M and H-H pairs regardless of pair distance, as in ordinary Jastrow factors. Here, we adopt a simple power-law decaying form of $\eta(r)$ as,

$$\eta(r) = 1 - \frac{a}{r^\kappa} \qquad (r \geq 1), \qquad (17)$$

where a and κ are variational parameters with ranges $0 \leq a \leq 1$ and $0 \leq \kappa \leq \infty$. Correspondingly, for $\mathcal{P}_Q$ in Ψ_{DH}, we employ the power-law decaying form $\Psi_{\mathrm{DH}}^{\mathrm{pwr}}$ in eq. (15b). Although the form of eq. (17) may not be the best, it is sufficient to ascertain the importance of the repulsive Jastrow factor for the Mott transition.

Incidentally, we compare the above wave functions with that studied in related papers by Capello *et al.*:[25]

$$\Psi = \exp\left[-\frac{1}{2} \sum_{i,j} v_{i,j}(n_i - 1)(n_j - 1) \right] \mathcal{P}_Q \Phi_0, \quad (18)$$

in which they emphasized importance of the long-range Jastrow factor. Equation (18) already has an attractive D-H factor in the Jastrow (exponential) part, but at each (i, j) the weight of attractive D-H correlation is inseparably connected with that of repulsive D-D (H-H) correlation. For $U \gtrsim U_c/2$, since the site occupation number is almost restricted to $0 \leq p \leq 2$, the number operator is written as,

$$n_j = 1 + d_j - h_j, \qquad (19)$$

using a doublon operator d_i in the space of $p \leq 2$:

$$d_i |i\rangle = \begin{cases} 1\, |i\rangle, & |i\rangle: \mathrm{D} \\ 0\, |i\rangle, & |i\rangle: \mathrm{otherwise} \end{cases}. \quad (20)$$

Thereby, the exponential part of eq. (18) becomes,

$$\exp\left[-\frac{1}{2} \sum_{i,j} v_{i,j}\,(d_i d_j + h_i h_j - d_i h_j - h_i d_j) \right], \quad (21)$$

which shows the attractive factor is a reciprocal of the repulsive factor. Thus, to adjust the D-H binding effect independently of the D-D correlation, one is obliged to add a redundant $\mathcal{P}_Q$ term. In this context, Ψ_J have the advantage in distinguishing the effect of D-H binding factors from that of repulsive ones, although eq. (18) and Ψ_J may work similarly.

We first optimize these trial functions, applying an optimization VMC scheme to systems with up to 1,600 particles ($L = 40$). With the optimal parameters obtained, we calculate the expectation values of relevant physical quantities, using a conventional VMC method. With this procedure, we can obtain accurate variational results in most cases, except for statistical errors. In Appendix, we briefly explain some details of the VMC calculations carried out in this paper.

3. Short-Range Doublon-Holon Factor

In this section, we study the short-range D-H binding wave function QWF, which exhibits typical properties of D-H-binding types of wave functions. In §3.1, we study the energy of QWF to find out a Mott critical behavior, by comparing with extreme cases. In §3.2, we discuss the site-occupation number versus U/t, in connection with experiments of a quantum gas microscope. In §3.3, the existence of Mott transition is corroborated and its properties are studied by various quantities.

3.1 Overall behavior of QWF's energy

To begin with, we briefly check the optimized energy of QWF, eqs. (8)-(13). In Fig. 2(a), the total energies per site, E/t, are compared between GWF and QWF of $\mathcal{P}_Q(\mu, \mu')$ for SQL. For small U/t (roughly $U < U_c/2 \sim 10t$), the improvement of E/t on GWF is very small, indicating the D-H correlation plays a minor role for weakly interacting conductive states. For $U \gtrsim U_c/2$, however, the curve of QWF gradually departs from that of GWF and approaches the extreme case of $\mathcal{P}_Q(1, 0)$, in which a doublon(s) and a holon(s) rigidly adhere to each other in the NN sites. Because such a state strongly suggests an insulator, a Mott transition is expected to occur near the value where the two curves join ($U/t \sim 20$). The fact that the energy of QWF is considerably lowered from that of GWF in this regime indicates the D-H binding effect play a prominent part in the Mott physics. The transition arising in GWF is a continuous type, because E/t vanishes as $\propto (1 - U/U_{\mathrm{BR}})^2$ in the Bose fluid side. In contrast, QWF exhibits a first-order transition, as we will see in §3.3. Hence, the mechanisms of the two transitions are qualitatively distinct.

The behavior in the regime of large U/t should obey an effective Hamiltonian of eq. (1) in the limit of $t/U \to 0$,

which is led by a canonical transformation[40] e^{iS} with

$$S = \frac{it}{U} \sum_{\langle i,j \rangle} \left(d_i b_i^\dagger b_j s_j + d_j b_j^\dagger b_i s_i - s_i b_i^\dagger b_j d_j - s_j b_j^\dagger b_i d_i \right),$$

(22)

where s_i is a projection operator of a singly-occupied site:

$$s_i |i\rangle = \begin{cases} 1 \, |i\rangle, & |i\rangle : \text{singly occupied} \\ 0 \, |i\rangle, & |i\rangle : \text{otherwise} \end{cases}.$$

(23)

Then, we have an expression at $n = 1$ in the space without D and H as,

$$H_{\text{eff}} = -\frac{4t^2}{U} \sum_{\langle i,j \rangle} s_i s_j = -2z N_s \frac{t^2}{U},$$

(24)

which is drawn as dash-dotted lines in Figs. 2(a) and 2(b). The energy of QWF well coincides with eq. (24) for a wide range of U ($> U_c$), meaning density fluctuation is properly introduced in the insulating regime. In Fig. 2(b), the same quantity is plotted for TAL. Because the behavior is qualitatively identical, henceforth, we address only SQL in most cases.
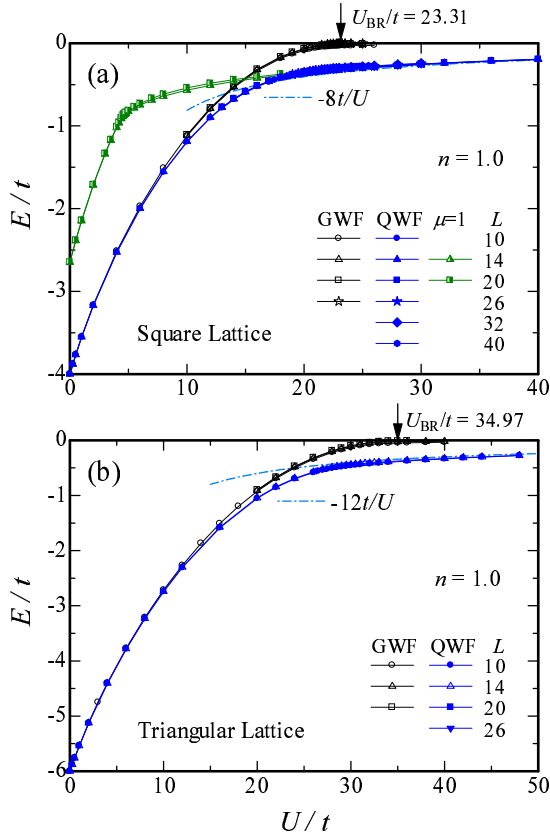

Fig. 2. (Color online) Energy expectation values of GWF (open symbols) and QWF (solid symbols) as a function of correlation strength U/t for (a) the square and (b) the triangular lattices. The result of strong-coupling expansion, eq. (24), is drawn by the dash-dotted lines. The critical values of Brinkman-Rice-type transitions in GWF are indicated by arrows on the upper axes. In (a), the data of QWF for $\mathcal{P}_Q(\mu = 1, \mu' = 0)$ are plotted with half-solid symbols (expressed as "$\mu = 1$"). The system-size dependence is inconspicuous in this scale.

3.2 Site-occupation number and parity correlation

In connection with the Mott transition, a change in the distribution [$P(p)$] of site-occupation number p have been directly observed in recent experiments of cold bosonic atom gases.[41, 42] Since noninteracting bosons are randomly distributed to the lattice sites, the ratio of the number of sites occupied by p particles (N_p) at $U/t = 0$ should obey a Poisson distribution:

$$P(p) = \frac{N_p}{N_s} = e^{-n} \frac{n^p}{p!},$$

(25)

with $n = 1$ at unit filling. As U/t increases, however, N_p with large p rapidly decreases to reduce $\langle H_U \rangle = U \langle D \rangle$, and for $U > U_c$ a number squeezed state is considered to be realize, in which most sites are loaded with one particle, $P(1) \sim 1$. The U/t dependence of $P(p)$ has been addressed using GWF,[21] a mean field approximation[43] and QMC with analytic confirmations.[44] With these studies in mind, we discuss the results of Ψ_{DH}.

Figure 3(a) shows the evolution of $P(p)$ for QWF, as U/t is varied. The Poissonian is changed to a shape of a symmetric peak centered at $p = 1$ at a relatively small value of U/t, and $P(p)$ with $p \geq 3$ almost vanishes; at $U/t = 12$, $P(3)$ is less than $P(2)/100$ (see Table I). The U/t dependence of $P(p)$ depends very slightly on the type of D-H correlation factor, as shown in Table I, where the values of $P(p)$ are compared among various correlation
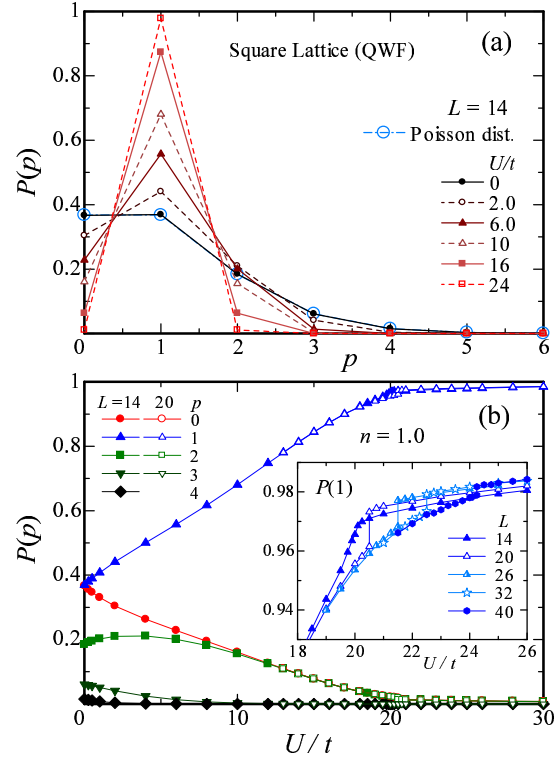

Fig. 3. (Color online) Ratio of the number of sites occupied by p particles (N_p/N_s) calculated with QWF are shown, (a) as a function of p for some U/t, and (b) as a function of U/t for small p. In (a) the Poisson distribution eq. (25) is plotted with large open circles to confirm the VMC outcome for $U/t = 0$. System-size dependence is negligible except for the Mott critical region, whose magnification for the case of $p = 1$ is shown in the inset in (b) for several L.

Table I.   Comparison of probability of site occupation $P(p)$ with $p = 0$-3 among GWF and four types of Ψ_{DH} for three values of U/t for $L = 20$. For $U/t \geq 12$, $P(4)$ is less than 10^{-5} in every case. In the state column, SF and MI denote 'superfluid' and 'Mott insulating', respectively. Incidentally, $P(0)$ and $P(2)$ of GWF must completely vanish for $L = \infty$ in the MI regime.

Ψ	state	$P(0)$	$P(1)$	$P(2)$	$P(3)$
$U/t = 0$ (Poisson distribution)					
-	SF	0.36788	0.36788	0.18394	0.06131
$U/t = 12$					
GWF	SF	0.14503	0.71147	0.14197	0.00153
QWF	SF	0.12694	0.74717	0.12485	0.00104
exp.	SF	0.13053	0.73986	0.12869	0.00092
pwr.	SF	0.12742	0.74605	0.12565	0.00089
opt.	SF	0.12670	0.74755	0.12481	0.00095
$U/t = 19$					
GWF	SF	0.04958	0.90084	0.04956	0.00001
QWF	SF	0.02982	0.94038	0.02979	0.00002
exp.	MI	0.01471	0.97057	0.01471	0.00000
pwr.	SF	0.02866	0.94269	0.02864	0.00001
opt.	SF	0.02855	0.94290	0.02854	0.00001
$U/t = 24$					
GWF	MI	0.00108	0.99783	0.00108	0.00000
QWF	MI	0.01016	0.97968	0.01016	0.00000
exp.	MI	0.01014	0.97973	0.01013	0.00000
pwr.	MI	0.01035	0.97931	0.01035	0.00000
opt.	MI	0.01038	0.97925	0.01038	0.00000

factors for some values of U/t. Furthermore, the U/t dependence shown in Fig. 3(b) is quantitatively consistent with a result of QMC [Fig. 1(b) in ref. 44]. Thus, in the Mott critical regime ($U \gtrsim U_{\mathrm{c}}/2$), we can safely consider the problem in the restricted space of $p \leq 2$, similarly to Fermi systems.

As seen in Fig. 3(b), the ratio of singly-occupied sites (doublon and holon densities) increases (decrease) almost linearly with U/t until near the critical point. The variation of $P(1)$ at the Mott critical point is as small as 1% in QWF [see the inset of Fig. 3(b)], and local number fluctuation remains in some degree even in the Mott insulating phase. Thus, a number squeezed state is gradually approached as U/t increases, and does not distinctly characterize a Mott insulating (MI) state.
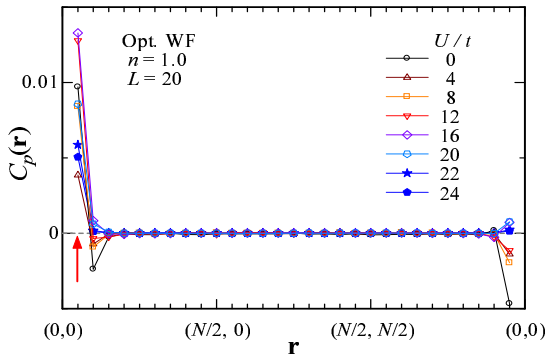

Fig. 4.   (Color online) Real-space parity correlation function for various values of U/t calculated with long-range D-H binding wave function $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ along path shown on abscissa. The arrow indicates the nearest-neighbor site. The data in the Mott insulating regime are indicated by solid symbols.

Recent development of single atom-single lattice site imaging technique (a quantum gas microscope) enables one to directly observe the parity (even or odd) of the occupied particle number p at each site;[45, 46] the parity of the site j is written as,

$$\bar{p}_j = \frac{1}{2}\left[1 - (-1)^{n_j}\right]. \qquad (26)$$

Thus, the correlation function of the parity,

$$C_{\bar{p}}(j,\ell) = \langle \bar{p}_j \bar{p}_\ell \rangle - \langle \bar{p}_j \rangle \langle \bar{p}_\ell \rangle \qquad (27)$$

becomes a quantity directly measured by experiments.[47] In Fig. 4, we show $C_{\bar{p}}(j,\ell)$ for various values of U/t using the long-range Ψ_{DH} with $\mathcal{P}_Q(f)$. Because $C_{\bar{p}}(j,\ell)$ is closely related to the usual density correlation function,

$$N(j,\ell) = \langle n_j n_\ell \rangle - \langle n_j \rangle \langle n_\ell \rangle, \qquad (28)$$

the magnitude of $C_{\bar{p}}(j,\ell)$ rapidly decays, as $|\mathbf{r}|$ increases. For large U/t, in particular in the insulating regime, the parity correlation is almost restricted to the nearest-neighbor sites, suggesting insignificance of the long-range part of correlation. In this regime, the parity operator eq. (26) is written by the projection operators eqs. (12) and (20) as,

$$\bar{p}_j = 1 - d_j - h_j, \qquad (29)$$

and n_j by eq. (19), so that $C_{\bar{p}}(j,\ell)$ is written with $N(j,\ell)$ and the doublon-holon correlation function:

$$C_{\mathrm{DH}}(j,\ell) = \langle d_j h_\ell \rangle - \langle d_j \rangle \langle h_\ell \rangle, \qquad (30)$$

as,

$$C_{\bar{p}}(j,\ell) = N(j,\ell) + 4C_{\mathrm{DH}}(j,\ell). \qquad (31)$$

Thus, the density and D-H correlation functions are fundamental also in analyzing quantum gas microscope experiments near the Mott transition.

3.3   Mott transition in QWF

Now, we analyze the superfluid-insulator transition in QWF more in detail. When we carefully watch the behavior of E/t in a magnified figure (Fig. 5), we notice a cusp for each L with $L \geq 20$. In each side of the cusp, we can find another local minimum (a metastable point) of E/t, which is smoothly extrapolated from the curve in the other side of the cusp, suggesting a first-order transition. This is supported by the existence of a discontinuity at the cusp point in the optimized variational parameters, as shown in Fig. 6. Critical values thus obtained are listed in Table III; the system-size dependence of U_{c}/t will be discussed in §4.2. A clear sign of a first-order transition is not observed for small systems ($L \leq 14$ in this case), similarly to fermionic systems.[32] For TAL, a clear discontinuity does not appear even for $L = 26$, as shown in Fig. 7 for μ, and E/t has very broad minimum at $U \sim U_{\mathrm{c}}$; the tendency toward a continuous transition is stronger for TAL, which tendency is similar to those of frustrated metallic states in fermionic models.[32]

In contrast to GWF, which has $g = 0$ for $U > U_{\mathrm{BR}}$, QWF has finite g even for $U > U_{\mathrm{c}}$, indicating the existence of density fluctuation even in the MI regime. The

nearest-neighbor D-H parameter μ exhibits a large discontinuity at U_c, and becomes close to 1 for $U > U_c$; the D-H binding becomes firm in the insulating regime. Note that, as we will see, the Mott transition is induced by the collaboration of the suppression of onsite density fluctuation by g and the D-H binding effect by $\mathcal{P}_Q$.

To see the properties of this first-order transition, we discuss some relevant quantities. We start with the expectation value of onsite correlation operator D, eq. (7). Especially at unit filling ($\langle n_j \rangle = \langle n_j \rangle^2 = 1$), this quantity

per site, $\tilde{d} = \langle D \rangle / N_s$, coincides with half of the onsite density fluctuation or variance,

$$\sigma_o^2 \equiv N(j,j) = \langle n_j^2 \rangle - \langle n_j \rangle^2 \qquad (32)$$

$$= \langle n_j^2 \rangle - \langle n_j \rangle = 2\tilde{d} \qquad (n=1). \qquad (33)$$

Substituting eq. (19) in eq. (33), we have a relation, $\tilde{d} = d$, valid for $U \gtrsim U_c/2$, where d is the doublon density. Thus, d is a key indicator of Mott transitions like for fermionic systems, so that we regard d as an order parameter of Mott transitions also for bosons. In Fig. 8, we show the U/t dependence of $\tilde{d}$ for GWF and QWF, and its magnification near U_c in the inset. For each value of L, $\tilde{d}$ of QWF decreases with U/t and exhibits a discontinuity at each U_c/t ($L \geq 20$). Conversely, the kinetic energy $E_t = \langle H_t \rangle$ increases with a discontinuity at U_c (not shown). This energetics meets the ordinary criterion of Mott transitions: The energy is stabilized by lowering E_U at the cost of E_t for $U > U_c$. The doublon density $\tilde{d}$ for QWF remains finite in the insulating regime, because local density fluctuation is permitted and transient
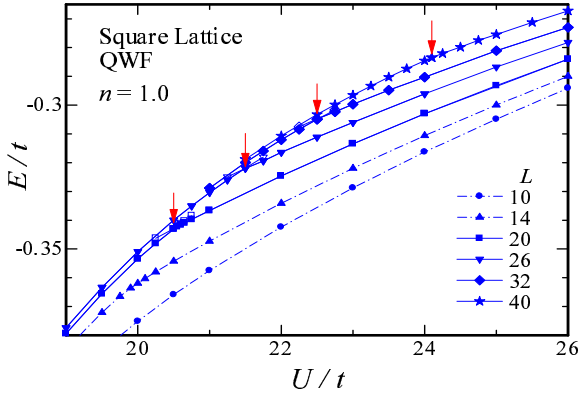

Fig. 5. (Color online) Magnification of total energy of QWF [Fig. 2(a)] near Mott critical points as function of interaction strength for several values of L. The arrows indicate the transition points (U_c/t). Open symbols near U_c/t denote the data of metastable states.
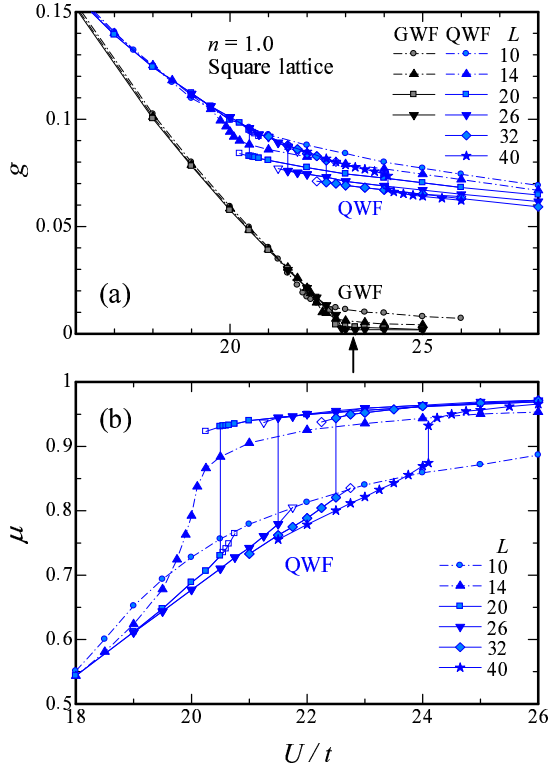

Fig. 6. (Color online) Optimized main variational parameters in QWF for square lattice, (a) Gutzwiller parameter and (b) nearest-neighbor D-H parameter, near Mott critical values for some L. In (a), corresponding values of GWF are added with an arrow indicating U_{BR}. Open symbols near U_c/t denote the values of metastable states.
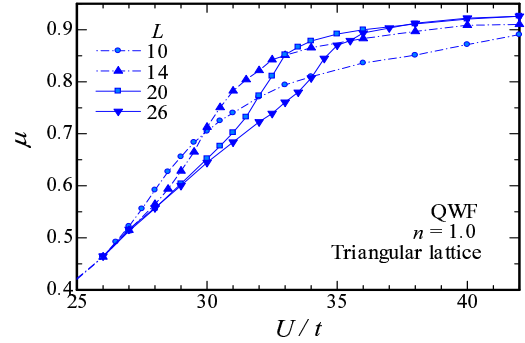

Fig. 7. (Color online) Optimized nearest-neighbor D-H parameter near Mott critical values in QWF for triangular lattice. Four system sizes (L) are compared.
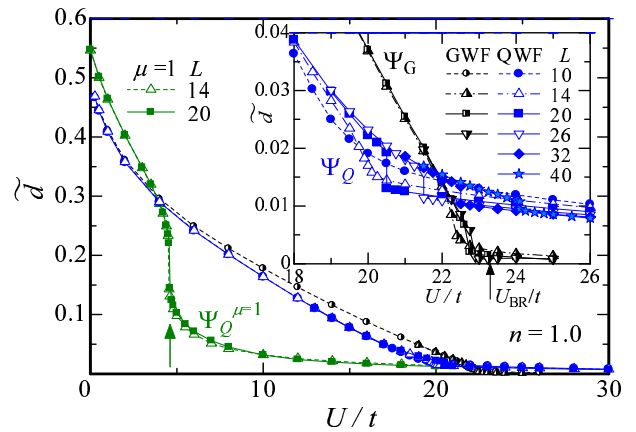

Fig. 8. (Color online) Comparison of expectation values of onsite correlation operator D [eq. (7)] per siteon the square lattice among GWF, QWF and completely D-H bound state $\Psi_Q^{\mu=1}$ (discussed later). This quantity is substantially doublon density d for $U/t \gtrsim 10$, and is half the onsite density fluctuation σ_o^2. The arrow on the curve of $\Psi_Q^{\mu=1}$ indicates the Mott critical point. The inset shows the magnification near the Mott critical points for GWF and QWF. The symbols in the inset are common to those in the main panel.

D-H pairs exist within neighboring sites [see also $P(2)$ in Table I]. This behavior contrasts with that of GWF, for which d vanishes for $U > U_{\mathrm{BR}}$. Note that this quantity will be measured as a function of U/t by experiments like quantum gas microscopes.
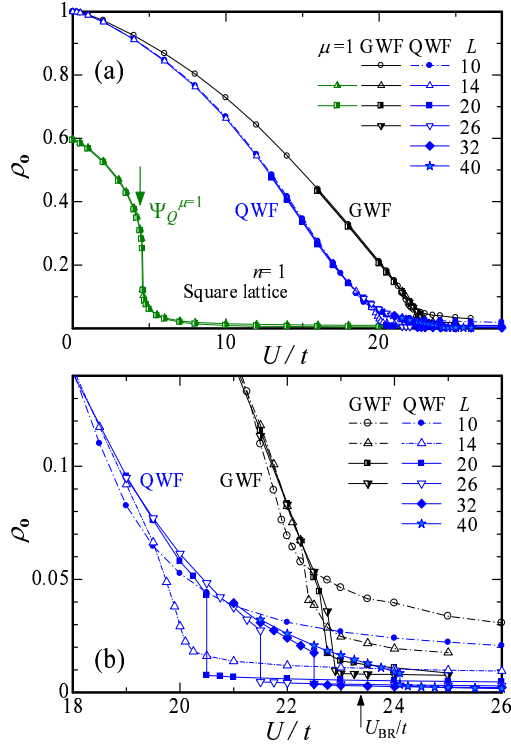

Fig. 9. (Color online) (a) Condensate fraction as a function of U/t for some values of L, calculated with three functions, GWF, QWF, and completely D-H bound function, $\Psi_Q^{\mu=1}$ (discussed later). The arrow on $\Psi_Q^{\mu=1}$ indicates the Mott critical point. (b) Magnification of the Mott critical region for GWF and QWF.

Next, we consider U/t dependence of the condensate fraction,

$$\rho_0 = n(\mathbf{0})/N_s, \tag{34}$$

with $n(\mathbf{0})$ being the $\mathbf{k} = \mathbf{0}$ element of the momentum distribution function:

$$n(\mathbf{k}) = \langle b_{\mathbf{k}}^\dagger b_{\mathbf{k}} \rangle = \frac{1}{N_s} \sum_{j,\ell} e^{i\mathbf{k}\cdot\mathbf{r}_\ell} \langle b_{j+\ell}^\dagger b_j \rangle. \tag{35}$$

In Fig. 9, we plot ρ_0 calculated with GWF and QWF. As U/t increases, the condensate fraction diminishes from the value of free bosons $\rho_0 = 1$, and vanishes at $U = U_c$ as the order of n/N_s in the non-superfluid states for all the wave functions. Thus, superfluidity vanishes at $U = U_c$. As shown in Fig. 9(b), the discontinuities appear for QWF with $L \geq 20$ in accordance with other quantities. Recently, the condensate fraction has been actually observed by experiments of cold atoms.[48,49]

Here, we discuss $\mathbf{k} \neq \mathbf{0}$ elements of $n(\mathbf{k})$, because it is a directly-observed quantity by cold-atom experiments. Figure 10 shows $n(\mathbf{k})$ obtained by the VMC calculations under some conditions. Because GWF gives the result identical with that of the Gutzwiller approximation in

fermionic cases,[23] $n(\mathbf{k})$ is constant for $\mathbf{k} \neq \mathbf{0}$ in both superfluid and insulating phases. On the other hand, wave functions with D-H factors yields dispersive $n(\mathbf{k})$, owing to the effect of density fluctuation. In Fig. 10, we draw $n(\mathbf{k})$ of the best D-H binding wave function in this study $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$. It is marked that the difference is small between the superfluid ($U/t = 20$) and insulating states ($U/t = 24$). As $\mathbf{k}$ leaves the $\Gamma(= \mathbf{0})$ point, $n(\mathbf{k})$ decreases in any direction, but its decrement is larger in the $\Gamma \to (\pi, \pi)$ direction than in the $\Gamma \to (\pi, 0)$ direction. This direction-dependent dispersion near the Mott critical point is actually observed as a cross-like intensity in absorption images of cold atoms, for example in Figs. 2f and 2g of ref. 2. This topic was previously argued by the perturbative correction to the Gutzwiller solution.[50]

Finally, to corroborate a superfluid-insulator transition at unit filling, we compare particle-density (n) dependence of various quantities in the vicinity of $n = 1$[51] between $U < U_c$ and $U > U_c$. Let us start with the condensate fraction ρ_0, which is plotted in Fig. 11(a). For $U = 16t$ ($< U_c$), ρ_0 becomes minimum at $n = 1$, but preserves a finite magnitude, whereas for $U/t = 24$ ($> U_c/t$), ρ_0 decreases as n approaches 1, and almost vanishes at $n = 1$. The sign of second derivative $\partial^2 \rho_0/\partial n^2$ becomes different between the two cases near unit filling. Thus, a MI state appears only at $n = 1$ and $U > U_c$. This is supported by the similar behavior of the onsite density fluctuation σ_o^2, eq. (32), as shown in the inset of Fig. 11(b). The order parameter of Mott transitions, $\tilde{d}$, depicted in the main panel of Fig. 11(b) monotonically increases for $U < U_c$, whereas it sharply decreases at $n = 1$ for $U > U_c$, suggesting a Mott transition. Figure 11(c) represents the chemical potential,

$$\zeta\left(N - \frac{1}{2}\right) = E(N) - E(N-1) = \frac{\partial E}{\partial n} \tag{36}$$

as a function of n. Although for $U < U_c$, ζ/t is always a smooth function of n, for $U > U_c$, it has a discontinuity corresponding to the Mott gap, broadly estimated as $\Delta/t = 8.2$ for $U/t = 24$ on SQL, and 17.6 for $U/t = 40$ on TAL (figure not shown). This is a direct evidence of
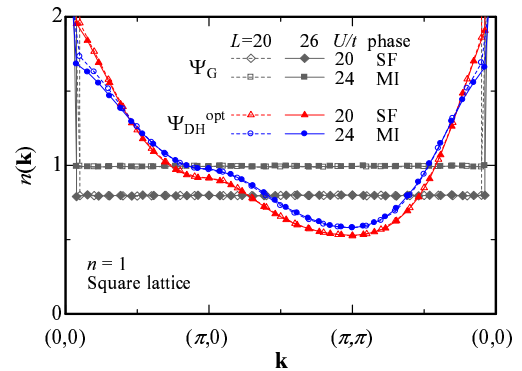

Fig. 10. (Color online) The momentum distribution function, eq. (35), are compared between GWF (Ψ_{G}) and optimizing wave function ($\Psi_{\mathrm{DH}}^{\mathrm{opt}}$), and between superfluid (SF) and Mott insulating (MI) phases. The value of the coherent wave number $\mathbf{k}=\mathbf{0}$ is omitted. Two system sizes are used. The behavior of $n(\mathbf{k})$ in the other D-H binding wave functions is similar to that of $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$.
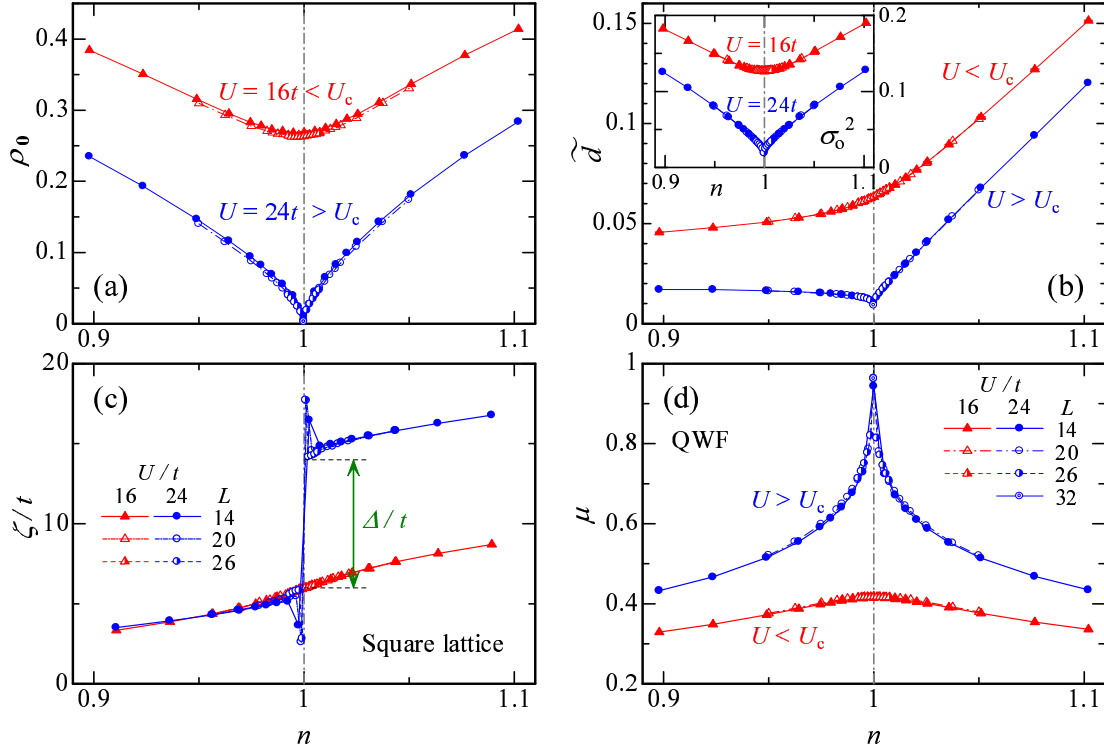
Fig. 11. (Color online) Particle-density dependence of various quantities are compared between at $U/t = 16$ [a superfluid (SF) state for $n = 1$], and at $U/t = 24$ [a Mott insulating (MI) state for $n = 1$], calculated with QWF on the square lattices of some system sizes $L \times L$. (a) Condensate fraction given by $n(\mathbf{0})/N_\mathrm{s}$. (b) The main panel shows the substantial doublon density, $\langle D \rangle / N_\mathrm{s}$, and the inset the onsite density fluctuation $\sigma_\mathrm{o}^2 = N(j,j)$ defined by eq. (32). For $n \neq 1$, the relation $\sigma_\mathrm{o}^2 = 2\tilde{d}$ [eqs. (32) and (33)] does not hold. (c) Chemical potential, eq. (36), actually estimated from the finite difference of E/t with respect to n. The arrow indicates the Mott gap for $U/t = 24$ at $n = 1$. (d) Optimized nearest-neighbor D-H binding parameter.

the Mott transition.

To recognize the importance of the D-H binding effect on the Mott transition, we compare, in Fig. 11(d), the optimized D-H binding parameter μ in QWF between the two phases. In the two phases, μ is symmetric with respect to $n = 1$, and has a maximum at $n = 1$, but the magnitude is distinct. For $U = 16t$, μ slowly varies with n and is still small at $n = 1$, whereas for $U = 24t$, μ anomalously increases as n approaches 1, and almost reaches 1 as L increases. Thus, the D-H binding effect is significantly enhanced in the very vicinity of the insulating state.

4. Long-Range Correlation Factors

In this section, we discuss the effect of long-range correlation factors, which we have disregarded in the preceding section. In §4.1 and §4.2, we focus on the properties of D-H attractive factors. In §4.3 we study the effect of additional D-D and H-H repulsive factors.

4.1 Optimized D-H attractive factors

To begin with, we compare the minimized energy among QWF and the wave functions with three long-range D-H attractive factors, exponentially decaying ($\Psi_\mathrm{DH}^\mathrm{exp}$), power-law decaying ($\Psi_\mathrm{DH}^\mathrm{pwr}$), and completely optimizing ($\Psi_\mathrm{DH}^\mathrm{opt}$) types, introduced in §2.2. As will be discussed in §4.2, a Mott transition occurs in each wave function near U_c of QWF ($\sim 20t$). For sufficiently large

and small U, compared with U_c, the total energies of the four functions are close to one another. As shown in Fig. 12, some difference appears in the Mott criti-
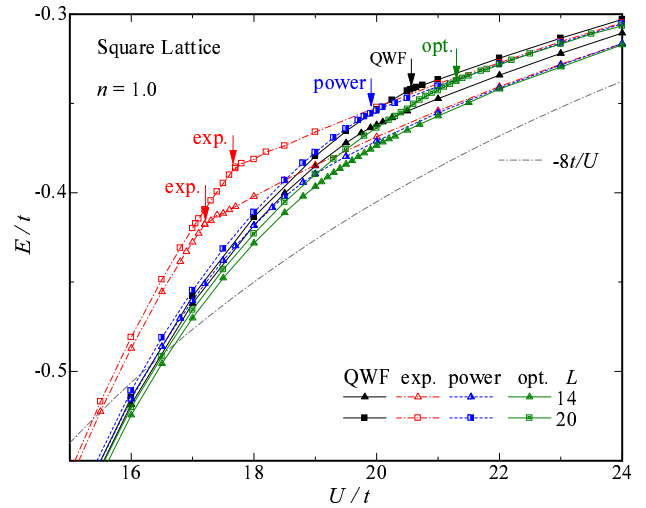

Fig. 12. (Color online) The total energies are compared near the Mott critical points among the wave functions with four kinds of D-H attractive factors: a short-ranged type (QWF) and three long-range types (exponentially- and power-law-decaying and completely-optimizing types). For clarity, only the data of two system sizes are plotted. The Mott critical point for each function ($L = 20$, and 14 for $\Psi_\mathrm{DH}^\mathrm{exp}$) are indicated by arrows. The value of strong-coupling expansion ($-8t/U$) is added.

cal region. Of course, E^{opt} (E for $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$) is always the lowest. For $U/t \gtrsim 21$, the energies of three long-range D-H wave functions are broadly analogous and somewhat improved over E^{QWF} for each system size. On the other hand, for $U/t \lesssim 20$, E^{exp} becomes clearly higher than E/t of the other D-H wave functions; E^{pwr} is also slightly higher than E^{QWF}. Now, we are aware that the frequently-used QWF yields a relatively good result, especially for $U < U_{\mathrm{c}}$, despite its simplicity. We analyze these features of E/t by comparing the optimized forms of correlation factors $f(r)$ among the D-H wave functions, in the following.
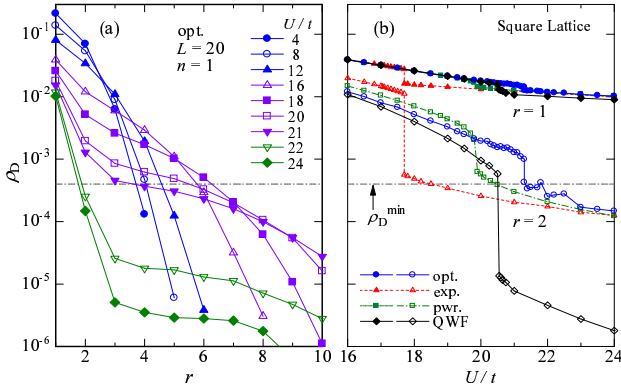

Fig. 13.   (Color online) The appearance probability of a doublon the distance from which to the nearest holon is r is plotted, (a) for the completely optimizing wave function $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ ($U_{\mathrm{c}}/t \sim 21.3$), versus r and for some values of U/t, and (b) for the four D-H wave functions versus U/t for $r = 1$ and 2. The threshold value $\rho_{\mathrm{D}}^{\mathrm{min}} = 0.0004$ is indicated by a dash-dotted line in both panels.

First, we discuss the effective range of $f(r)$. We represent the distance from a doublon (multiplon) to its nearest holon simply by r here, and the probability that a site is occupied by a doublon of r by $\rho_{\mathrm{D}}(r)$, which satisfies the relation,

$$\sum_{r=1}^{L} \rho_{\mathrm{D}}(r) = \tilde{d}. \tag{37}$$

In Fig. 13(a), we plot $\rho_{\mathrm{D}}(r)$ of $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ in a logarithmic scale. Because $\rho_{\mathrm{D}}(r)$ is a monotonically decreasing function of r, it is convenient to define a threshold value $\rho_{\mathrm{D}}^{\mathrm{min}}$, which broadly gives the effective range of $f(r)$ by,

$$\rho_{\mathrm{D}}(r) > \rho_{\mathrm{D}}^{\mathrm{min}}. \tag{38}$$

The contribution from $\rho_{\mathrm{D}}(r) < \rho_{\mathrm{D}}^{\mathrm{min}}$ should be negligible, namely, the corresponding particle configurations should appear very rarely. Thus, the optimized values of $f(r)$ for such distant r are unreliable and insignificant. Here, we set $\rho_{\mathrm{D}}^{\mathrm{min}} = 100/250,000$ on the basis of accuracy in the VMC calculations. In a weakly interacting regime ($U/t \lesssim 12$), the weight of $\rho_{\mathrm{D}}(r)$ concentrates in $r \lesssim 4$, and virtually vanishes ($\rho_{\mathrm{D}}(r) < \rho_{\mathrm{D}}^{\mathrm{min}}$) for larger r, because multiplons and holons are crowded. In the superfluid regime, $\rho_{\mathrm{D}}(r)$ comes to decrease slowly as U/t increases, and extend the valid range to $r \lesssim 6$ near U_{c}/t. Meanwhile, in the insulating regime, $\rho_{\mathrm{D}}(r)$ immediately decays to have substantial weight only on $r = 1$ or at most 2. This contrastive feature of the effective range clearly implies that the nature of D-H binding effect changes at U_{c}. This effective range of $f(r)$ is closely related to the D-H binding length, ξ_{dh}, introduced in §5. Making a similar analysis with the condition (38) for the other wave functions, we determine the effective range of $f(r)$ for various values of U/t for each function.

Table II.   Comparison of appearance probability and weight of D-H factors for $r = 1$-3 among four wave functions in insulating regime ($U/t = 22$) for $L = 20$.

Ψ_{DH}	$\rho_{\mathrm{D}}(1)$ $\times 10^{-2}$	$\rho_{\mathrm{D}}(2)$ $\times 10^{-4}$	$f(2)$	$\rho_{\mathrm{D}}(3)$ $\times 10^{-5}$	$f(3)$	E/t
optim.	1.18	2.54	0.263	2.57	0.090	-0.3282
power	1.17	2.08	0.247	3.80	0.109	-0.3279
exp.	1.15	1.75	0.236	0.84	0.056	-0.3272
QWF	1.05	0.052	0.050	0.64	0.050	-0.3246

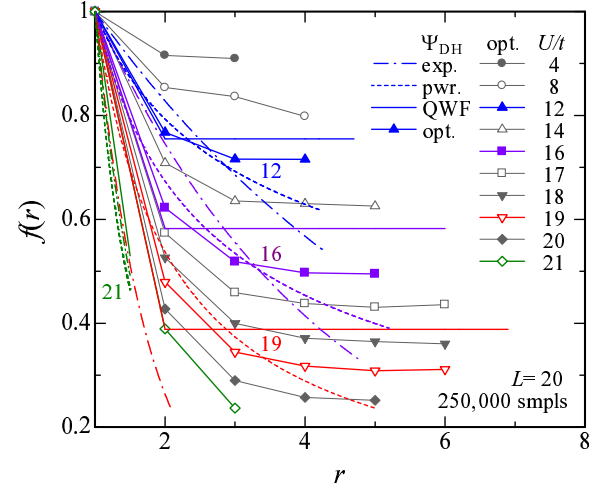
Fig. 14.   (Color online) The optimized D-H attractive factors are compared among four wave functions versus nearest D-to-H distance r. Except for $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$, the data for $U/t = 12$, 16, 19 and 21 are shown for clarity. The range of r is restricted according to the condition (38). The data of $U/t = 19$ for $\Psi_{\mathrm{DH}}^{\mathrm{exp}}$ and of $U/t = 21$ except for $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ are in the insulating phase. For QWF, we disregard the contribution of μ'.

Figure 14 compares the behavior of optimized $f(r)$ among the four wave functions within the effective range thus determined. In the conductive regime, $f(r)$ in $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ rapidly decreases for $r \lesssim 3$, but becomes almost constant for $r \gtrsim 3$. Namely, the D-H binding is effective only for $r \lesssim 3$, and a doublon is released from the bondage of holons for $r \gtrsim 3$. Because $f(r)$ of QWF is constant for $r \geq 3$, the behavior for large r is analogous to $f(r)$ in $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$. In contrast, $f(r)$ in $\Psi_{\mathrm{DH}}^{\mathrm{pwr}}$ and especially in $\Psi_{\mathrm{DH}}^{\mathrm{exp}}$ continues decreasing to zero as r increases, namely, the effective range of D-H binding is too long, compared with $f(r)$ in $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$. This is a cause of the unexpected good (unsatisfactory) result of QWF ($\Psi_{\mathrm{DH}}^{\mathrm{exp}}$). In the MI regime, $f(r)$ in $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ rapidly decays with r, and the effective range is limited to at most $r = 2$, as mentioned [Fig. 13(a)].

Regarding energy improvement, the appearance probability of nearest-neighbor D-H pairs seems primarily important; as listed in Table II, $\rho_D(1)$ is more than 10% smaller in QWF than in the long-range wave functions. In addition, the role of $f(2)$ is not negligible. In Fig. 13(b), $\rho_D(r)$ for $r = 1$ and 2 are compared in a logarithmic scale. For $U > U_c$, the three long-range wave functions have similar values of $\rho_D(2)$, whereas the values of QWF are roughly two orders of magnitude smaller. This is directly reflected in $f(2)$, as shown in Table II. No such great differences in $\rho_D(r)$ and $f(r)$ can be seen for $r \geq 3$ among the four wave functions [see also Fig. 15(b)].
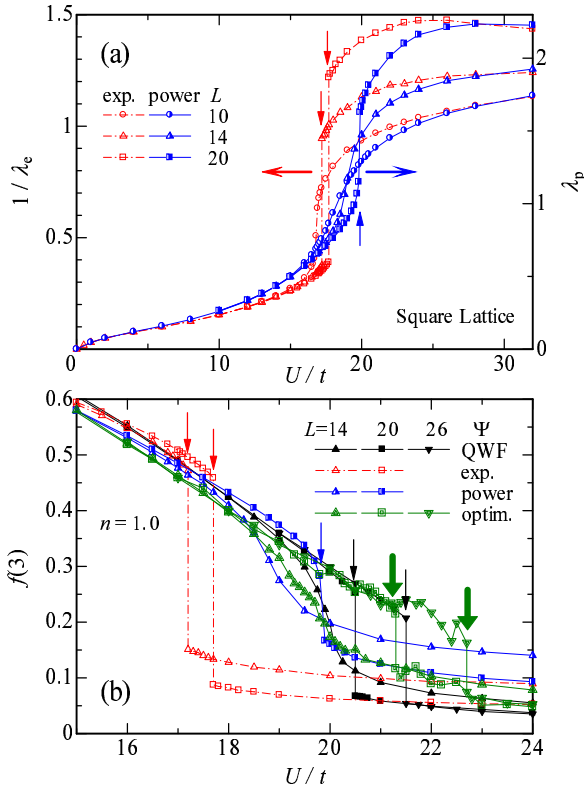
4.2   Mott transitions in D-H attractive factors

Fig. 15.   (Color online) (a) The optimized parameters controlling D-H correlation length, $1/\lambda_e$ in $\Psi_{\mathrm{DH}}^{\mathrm{exp}}$ and λ_p in $\Psi_{\mathrm{DH}}^{\mathrm{pwr}}$ are plotted for $L = 10\text{-}20$. (b) The weight of D-H attractive factors $f(r)$ for $r = 3$ is compared near the Mott critical points among the four D-H binding wave functions. In both panels, the critical points are indicated by vertical arrows, in case they are manifest. The thick arrows for $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ in (b) means large ambiguity in U_c/t.

In this subsection, we consider other properties of long-range wave functions as to the Mott transition.

We start with the parameters which controls the range of D-H attractive correlation, namely λ_e in $\Psi_{\mathrm{DH}}^{\mathrm{exp}}$ and λ_p in $\Psi_{\mathrm{DH}}^{\mathrm{pwr}}$ (Fig. 1). Their optimized values are shown in Fig. 15(a). Both λ_e and $1/\lambda_p$ abruptly increases at $U/t = 17\text{-}20$, namely, the D-H correlation range becomes short. In particular, clear jumps exist for $L \geq 14$ in $1/\lambda_e$ and $L = 20$ in λ_p, as indicated by arrows. In Fig. 15(b), the optimized D-H correlation weight $f(r)$ for $r = 3$ is magnified near the critical points. Both $\Psi_{\mathrm{DH}}^{\mathrm{exp}}$ and $\Psi_{\mathrm{DH}}^{\mathrm{pwr}}$

exhibit critical behavior at the same U_c/t as for λ; Jumps in $f(3)$ exist also for $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ with $L \geq 20$.
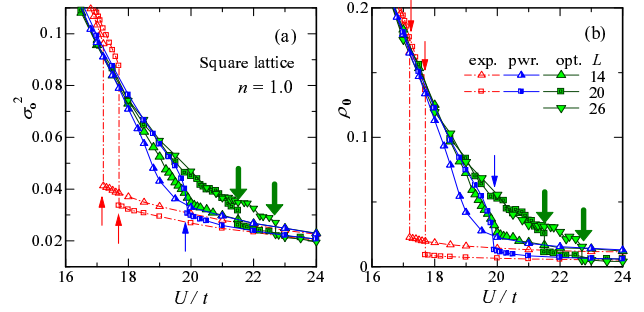

Fig. 16.   (Color online) (a) Variance of onsite density correlation function (or $2\tilde{d}$) and (b) condensate fraction, both calculated using wave functions with three kinds of long-range D-H factors, $\Psi_{\mathrm{DH}}^{\mathrm{exp}}$, $\Psi_{\mathrm{DH}}^{\mathrm{pwr}}$ and $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ for a few values of L. The arrows shows the Mott critical points.

Next, in Fig. 16, we show two quantities characterizing Mott transitions, σ_o [eq. (32)] or $2\tilde{d}$ [eq. (33)], and ρ_o [eq. (34)]. Again, both exhibit anomalous behavior at U_c/t determined above by the anomalies of variational parameters. The behavior is similar to that of QWF (Figs. 8 and 9). Thus, a first-order Mott transition can be described by a wide class of D-H binding wave functions if sufficiently large systems are considered.

Details of the critical behavior as well as the value of Mott critical point are different for different wave functions in some degree. As clear cusps appear in E^{exp} (Fig. 12), $\Psi_{\mathrm{DH}}^{\mathrm{exp}}$ exhibits sharp first-order critical behavior even for a small system of $L = 14$. Near U_c, two energy minima are distinguished, and a hysteresis is confirmed in the E/t-U/t plane. A hysteresis is also observed for $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ widely near U_c/t, and the energy difference between the two phases is very small. Thus, the control of optimization process becomes considerably difficult in the critical regime, so that the estimation of U_c/t for $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ is less accurate than for the others. This is probably caused by the redundancy of parameters. In comparison, $\Psi_{\mathrm{DH}}^{\mathrm{pwr}}$ behaves mildly near U_c and we have not detect a manifest hysteresis even for $L = 20$.

Finally, we discuss the value of the Mott critical point. In Table III, the critical values determined by the D-H binding wave functions in this work are listed with those of other studies. Among the three long-range Ψ_{DH} treated here, U_c/t is mutually different to some extent. On the basis of E/t in Fig. 12, this difference is considered to stem from the propriety of Ψ_{DH} in describing the superfluid state near U_c/t. Because E/t of the three Ψ_{DH}'s is similar to one another for $U > U_c$, a lower E/t or a better Ψ_{DH} in the conductive side of U_c yields a larger U_c/t. Furthermore, the critical value U_c/t, estimated by any Ψ_{DH} including QWF (§3.3), steadily increases as L increases. This is mainly because the system-size dependence of E/t is considerably large in the insulating side of U_c/t, but small in the conductive side, as in Figs. 5 and 12. These aspects of the critical value are contrary to reliable estimates by QMC and a strong-coupling

Table III. The Mott critical values U_c/t are compared among the wave functions treated in this paper for some system sizes. Sections with hyphens indicate the cases in which clear behavior of first-order transition is not observed. For empty sections, we have not performed calculations in this study. The last low shows the result, when the repulsive Jastrow factor is included (see §4.3). In the lower lines, the values obtained in other studies are entered.

Ψ (or method)	U_c/t $(n = 1,$ square lattice)					
$L \rightarrow$	10	14	20	26	32	40
QWF (Ψ_Q)	-	-	20.5	21.5	22.5	24.1
optim. (Ψ_{DH}^{opt})	-	-	~ 21.3	~ 22.7		
exp. (Ψ_{DH}^{exp})	-	17.2	17.65			
power (Ψ_{DH}^{pwr})	-	-	19.85			
repulsive (Ψ_J)	-	-	19.85			
GWF (U_{BR}/t)	$23.31\cdots$ (refs. 21 and 22)					
VMC (Jastrow)	~ 20.6 (ref. 25)					
Strong coupling	16.7 (ref. 16)					
Recent QMC	16.25 (ref. 14), 16.7 (ref. 15)					

Table IV. Comparison of total energy per site E/t near U_c/t among GWF and wave functions with power-law decaying type D-H and D-D correlation factors. The digits in the brackets indicate the ratios with respect to the values of GWF+D-H.

U/t	GWF	GWF+D-H	GWF+D-H+D-D
$L = 14$			
19	-0.11712 (0.301)	-0.38945	-0.39156 (1.005)
20	-0.07435 (0.200)	-0.37103	-0.37561 (1.012)
21	-0.04314 (0.121)	-0.35523	-0.35680 (1.004)
$L = 20$ $(U_c/t = 19.85)$			
19	-0.11042 (0.293)	-0.37742	-0.37858 (1.003)
20	-0.06750 (0.191)	-0.35410	-0.35405 (1.000)
21	-0.03612 (0.106)	-0.34034	-0.34138 (1.003)

expansion, $U_c/t = 16$-17. To obtain a more accurate critical value, some factor overlooked in the present class of wave functions may be taken into account. As a possibility, we take up D-D repulsive correlations in §4.3.

4.3   Effect of repulsive Jastrow factors

So far, we have disregarded the effect of intersite repulsive factors, because it is known long-range repulsive factors reduce the energy only slightly for fermions at half filling.[29,34] Here, we check that the Bose Hubbard model also has this property, and a repulsive factor is insufficient to improve the system-size dependence of U_c/t. For simplicity, we employ the power-law-decaying type both for attractive [eqs. (14) and (15b)] and repulsive [eqs. (16) and (17)] correlation factors mentioned in §2.

First, we argue improvement in energy by D-H attractive and D-D repulsive factors over GWF. Table IV lists the numerical values of E/t for the three wave functions and two system sizes. As U approaches U_{BR} (= 23.31t), E^{GWF} steadily increases toward zero. By introducing the D-H binding factor, E is significantly improved on E^{GWF} (70-90%), and comes to increase slowly as U/t increases, like the case of QWF in Fig. 2(a). On the other hand, when we add the D-D and H-H repulsive correlation to Ψ_{DH}, the improvement of E/t on Ψ_{DH} is very slight (mostly less than 1%) in both conductive and insulating

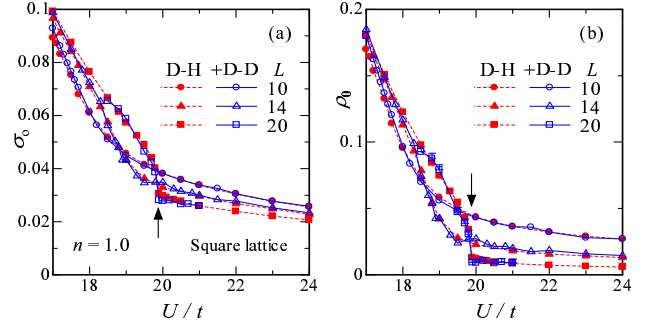phases, and comparable to the statistical errors, which are large for Ψ_J.[53]

Fig. 17.   (Color online) (a) The variance of onsite density correlation function (or $2\tilde{d}$) and (b) condensate fraction, are compared between Ψ_{DH}^{pwr} (D-H) and Ψ_J^{pwr} (+D-D) for three values of L. The arrows indicate U_c/t for $L = 20$.

Although the energy reduction is slight, we should confirm whether the D-D repulsive factor $\mathcal{P}_J$ improves physical quantities relevant to the Mott transition. Since $\mathcal{P}_J$ tends to lengthen the inter-doublon (D-D) distance in the insulating side (not shown), we have expected that Ψ_J^{pwr} lowers U_c/t. In Figs. 17(a) and Figs. 17(b), we compare σ_o and ρ_0, respectively, between Ψ_{DH}^{pwr} and Ψ_J^{pwr}. Against our anticipation, the differences in both quantities are negligibly small, and a meaningful shift of U_c/t cannot be observed (Table III).

Here, we have focused on the power-law decaying case. For fermions, it is found that various repulsive factors somewhat lower the values of U_c/t, but the situation in system-size dependence does not change.[34] Hence, we conclude that the effect of repulsive Jastrow factor is not significant, as far as the Mott transition is concerned. We will take no account of this factor in the following.

5.   Renewed Picture of Mott Transition

In a previous paper for electrons,[32] we affirmed a simple mechanism of Mott transitions, in which the binding of a doublon (plus charge carrier) to a holon (minus charge carrier) in the insulating regime and the release from binding in the conductive regime are the essence of the Mott transition. In this section, we extend this picture to comprehend a case of completely D-H binding.

First, we briefly mention the properties of the completely D-H bound state in NN sites, $\Psi_Q^{\mu=1}$, namely QWF with $\mu(= 1)$ and $\mu'(= 0)$ in eq. (9). Because, in $\Psi_Q^{\mu=1}$, a doublon must be accompanied by at least one holon in its four NN sites, we tend to regard it as insulating for any value of U/t. In fact, however, $\Psi_Q^{\mu=1}$ is a superfluid state for small U/t, and exhibit a Mott transition at $U/t = 4.55$. A sign of a Mott transition can be recognized in the cusp behavior of E/t at $U/t \sim 4.55$ in Fig. 2(a), and especially in sudden drop of the condensate fraction in Fig. 9(a). In Fig. 18, we show the optimized value of g and multiplon density $\tilde{d}$, besides the above two quantities. The discontinuous decreases of the order parameter $\tilde{d}$, and the sudden vanishing of ρ_0 corroborate the Mott transition at $U_c/t = 4.55$.

This Mott transition in $\Psi_Q^{\mu=1}$, especially in the conductive side, cannot be understood by a simple release from the D-H binding and mutually independent movement of the two kinds of carriers, because a doublon is in contact with a holon even in the conductive regime. Instead of this simple point of view, we propose an improved picture extensively applicable to Mott transitions. To this end, it is convenient to introduce two relevant length scales, D-H binding length ξ_{dh}, and minimum D-D (H-H) distance ξ_{dd}, which are loosely defined as follows: The distances of most of the nearest D-H pairs are smaller than ξ_{dh}; in other words, a doublon is seldom distant from a holon beyond ξ_{dh}. ξ_{dd} is a D-D (H-H) exclusion distance, namely, the inter-doublon (inter-holon) distances within which two doublons (or holons) are mutually almost inaccessible. In general, ξ_{dh} as well as ξ_{dd} depends on U/t.

Then, we postulate that an attractive correlation factor $\mathcal{P}_Q$ yields D-H pairs of a binding length ξ_{dh} according to U/t. An improved picture of Mott transition is schematically represented in Fig. 19. In the insulating phase, the relation $\xi_{\mathrm{dh}} < \xi_{\mathrm{dd}}$ holds, indicating that the domains of D-H pairs do not usually overlap, at least, not in sequence. As a result, most D-H pairs are isolated and a doublon and a holon are confined within ξ_{dh}, resulting in only local density fluctuation. To this point, the picture is basically identical with one previously proposed. In the conductive phase ($U < U_c$), ξ_{dh} becomes longer than ξ_{dd}, indicating the domains of D-H pairs overlap with one another. Then, a doublon in a D-H pair can exchange a partner holon with a holon in an adjacent D-H pair. Consequently, a doublon and a holon can move independently as carriers by exchanging the partner, as shown in long arrows in Fig. 19. It follows that, as the value of U/t is varied, a Mott transition takes place when ξ_{dh} becomes equivalent to ξ_{dd}, which (correctly ℓ_{dd}) is roughly $1/\sqrt{\tilde{d}}$, and is expected to be a monotonically increasing function of U/t.

To justify the above picture, we have to appropriately estimate ξ_{dh} and ξ_{dd}, and confirm that the Mott tran-
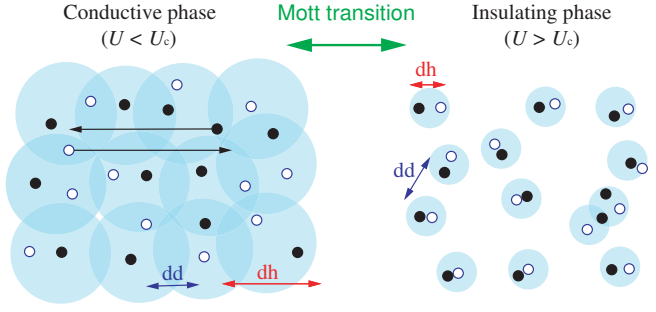

Fig. 19. (Color online) Schematic figure of improved picture of Mott transitions. Small solid and open circles denote doublons and holons, respectively. The D-D distance is closely related to ξ_{dd}, shown by "dd". Each pale circle indicates the domain of a D-H pair, whose diameter roughly corresponds to ξ_{dh}, shown by "dh". Typical configurations are drawn for the two phases. In each pair domain, the existence of at least one doublon and one holon is required, but excess doublons and holons can move around independently slipping out of the original domain, as indicated by long arrows for the conductive phase.

sitions take place when these two length scales intersect each other at U_c/t. After checking various cases, we have found that the following formulae work well and are physically natural to the above point of view,

$$\xi_{\mathrm{dh}} = \ell_{\mathrm{dh}} + \sigma_{\mathrm{dh}}, \tag{39}$$

$$\xi_{\mathrm{dd}} = \ell_{\mathrm{dd}} - \sigma_{\mathrm{dd}}, \tag{40}$$

where ℓ_{dh} (ℓ_{dd}) denotes the average of the nearest D-to-H and H-to-D (D-to-D and H-to-H) distance, and σ_Λ the standard deviation of ℓ_Λ, with Λ being an index of "dh" and "dd":

$$\sigma_\Lambda = \sqrt{\frac{1}{M} \sum_{m=1}^{M} (\ell_\Lambda^m - \ell_\Lambda)^2}. \tag{41}$$

Here, m runs over all the doublons and holons in all the measured samples, M indicates the total number of doublons and holons in all the measured samples, and ℓ_Λ^m indicates the nearest Λ distance for the m-th doublon (or holon). The addition and subtraction of σ in eqs. (39) and (40) represent a U/t-dependent "softness" of the D-H binding and D-D repulsive correlations, respectively.

First, we discuss a special case of the tightly D-H bound state $\Psi_Q^{\mu=1}$, shown in Fig. 20(a). Because ℓ_{dh} is restricted to 1 in $\Psi_Q^{\mu=1}$, $\sigma_{\mathrm{dh}} = 0$ and $\xi_{\mathrm{dh}} = 1$ hold, irrespective of the value of U/t. On the other hand, because the doublon density decreases as U/t increases, both ℓ_{dd} and σ_{dd} monotonically increase qualitatively in a manner similar to those of QWF shown in Fig. 21, and consequently, ξ_{dd} also monotonically increases with a discontinuity at U_c/t. As shown in Fig. 20(a), ξ_{dd} intersects ξ_{dh} in this discontinuity; this special case is consistent with the above picture.

Next, we consider more ordinary D-H binding wave functions: QWF, $\Psi_{\mathrm{DH}}^{\mathrm{exp}}$, $\Psi_{\mathrm{DH}}^{\mathrm{pwr}}$ and $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$. In Fig. 21, we
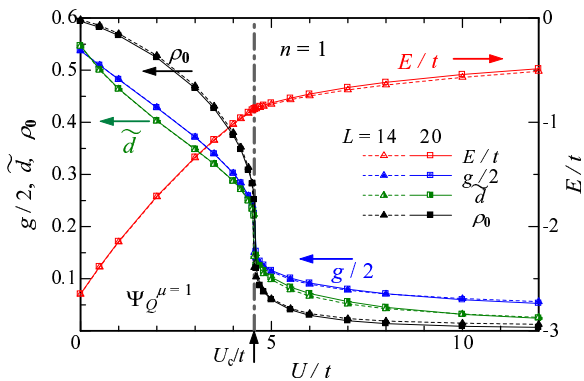

Fig. 18. (Color online) Four quantities calculated with the completely D-H bound state $\Psi_Q^{\mu=1}$ are plotted as a function of U/t; namely, total energy per site E/t (right axis), half of the optimized Gutzwiller parameter g, multiplon density $\tilde{d}$ and condensate fraction ρ_0. The Mott critical point U_c/t is indicated by a shadowed dash-dotted line. The system-size dependence is negligible.
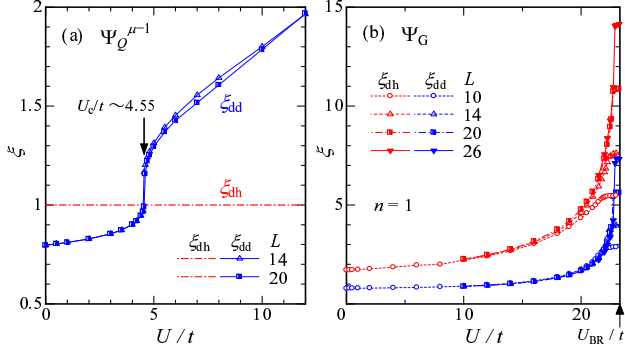
Fig. 20. (Color online) The D-H binding length ξ_{dh} and minimum D-D distance ξ_{dd} defined by eqs. (39) and (40) are plotted versus U/t for (a) the completely D-H bound state, and (b) the Gutzwiller wave function, for some values of L.
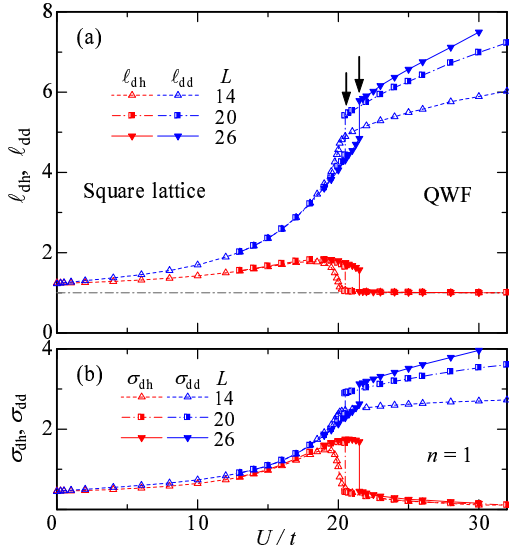

Fig. 21. (Color online) (a) The average nearest D-H and D-D distances and (b) their standard deviations are plotted for the short-range D-H binding wave function QWF as a function of U/t. The Mott critical values for $L = 20$ and 26 are indicated by arrows in (a). The scales of coordinates in (a) and (b) are common.

show the average nearest D-H and D-D distances and their standard deviations for QWF. The behavior of ℓ's and of σ's for the three long-range D-H binding wave functions is basically the same. As U/t increases, ℓ_{dd} rapidly increases, exhibits a discontinuity at U_{c}/t, and further increases for $U > U_{\mathrm{c}}$ with a considerable system-size dependence. Meanwhile, ℓ_{dh} at first gradually increases, and has a maximum in the vicinity of U_{c}/t, then decreases discontinuously at U_{c}/t, and converges to $\ell_{\mathrm{dh}} = 1$ for $U > U_{\mathrm{c}}$. The magnitude of σ_{dh} becomes large near U_{c}/t in the conductive side, but tends to vanish for $U > U_{\mathrm{c}}$. Substituting such ℓ and σ for eqs. (39) and (40), we obtain ξ_{dh} and ξ_{dd} for each of the four D-H binding states, and plot them in Fig.22. As expected from Fig. 21, ξ_{dd} is a monotonically increasing function
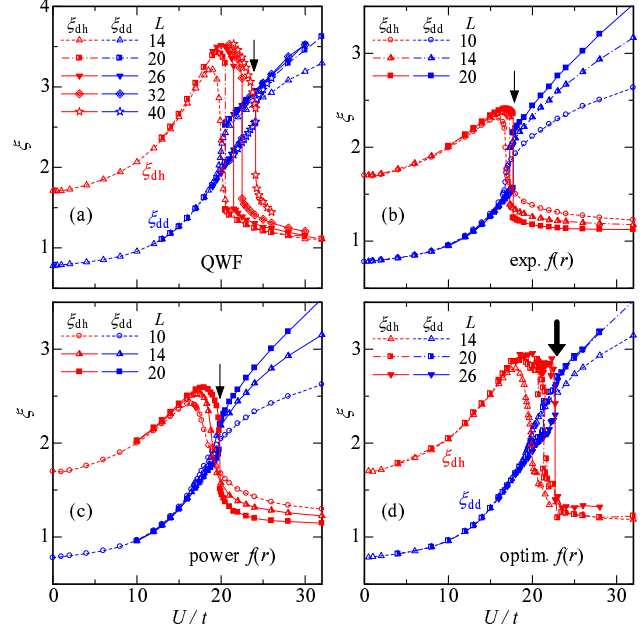

Fig. 22. (Color online) The D-H binding length ξ_{dh} and minimum D-D distance ξ_{dd} defined by eqs. (39) and (40) are plotted versus U/t for (a) the short-range D-H binding state QWF, (b) the exponentially-decaying D-H binding state $\Psi_{\mathrm{DH}}^{\mathrm{exp}}$, (c) the power-law-decaying D-H binding state $\Psi_{\mathrm{DH}}^{\mathrm{pwr}}$, (d) the optimized D-H binding state $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$, for some values of L. An arrow in each panel denotes the Mott critical point for the largest system size in each state.

of U/t with a discontinuity at U_{c}, whereas ξ_{dh} has a maximum just below U_{c} and becomes a decreasing function of U/t with a discontinuous drop at U_{c}. Consequently, the magnitude of the two length scales is reversed suddenly at U_{c}, and the relation $\xi_{\mathrm{dd}} > \xi_{\mathrm{dh}}$ holds for $U > U_{\mathrm{c}}$ for each L of every wave function. Thus, we conclude that a picture explained in Fig. 19 seems appropriate to the D-H binding mechanism of Mott transitions.

Incidentally, the above point of view is out of focus for a non-D-H binding mechanism like Brinkman-Rice-type transitions. As depicted in Fig. 20(b), for GWF, ξ_{dh} is always longer than ξ_{dd}, and does not show a tendency to decrease, up to the Brinkman-Rice point. The value of ξ at U_{BR}/t is proportional to L, which means that Ψ_{G} entirely lacks intersite correlations.

6. Summary

In this paper, we have studied the spinless Bose Hubbard models at unit filling on the square and triangular lattices, using a variational Monte Carlo scheme. Our primary aim is to grasp fundamentals of the Mott transition without influence of the spin degree of freedom. In the trial wave functions, we allow for various types of doublon-holon attractive factors and a doublon-doublon (holon-holon) repulsive factor in addition to the onsite repulsive (Gutzwiller) factor. We itemize the main results below.

(1) Because multiply occupied sites with more than two particles almost vanish for $U \gtrsim U_{\mathrm{c}}/2$ for all the wave

functions we check, we can consider the Mott physics near U_c/t (~ 20 for the square lattice) only with a doubly occupied site (doublon, D), an empty site (holon, H) and a singly occupied site, like fermionic cases.

(2) Wave functions with appropriate D-H attractive correlation factors exhibit first-order Mott transitions. Unlike the Brinkman-Rice-type transition arising in the Gutzwiller wave function, these transitions have density fluctuation in the Mott insulator side. By density fluctuation, anisotropy is introduced into the momentum distribution function, which is consistent with absorption images observed in cold atom experiments.[2] The total energy for $U > U_c$ coincides well with the results of the strong coupling expansion ($E/t \propto -t/U$).

(3) In the conductive (superfluid) state, the optimized value of D-H attractive correlation weight $f(r)$ [eq. (14)], with r being the interparticle distance, rapidly decreases with r for $r \lesssim 3$, and is almost constant for larger r. Thereby, we found it reasonable that the wave function with short-range D-H factors (QWF), in which $f(r)$ for $r \geq 3$ is constant, gives unexpectedly good results. On the other hand, in the Mott insulating state, $f(r)$ almost vanishes for $r \geq 3$, suggesting that a doublon and a holon confine each other in near-neighbor ($r \leq 2$) sites as a D-H pair. This will be confirmed by recently developed quantum gas microscope experiments.[45, 46]

(4) The critical behavior in the triangular lattice is more continuous-transition-like than in the square lattice. A similar phenomenon emerges in paramagnetic states of electronic systems,[32] and was ascribed to the frustration of spins. However, it is probable that its origin is closely related to the connectivity of lattices.

(5) We have improved the D-H binding picture of Mott transitions, by introducing two characteristic length scales, D-H binding length ξ_{dh}, which broadly represents the size of a D-H pair, and minimum D-D distance ξ_{dd}. We appropriately determined ξ_{dh} and ξ_{dd}. In the conductive state ($\xi_{\mathrm{dh}} > \xi_{\mathrm{dd}}$), because the domains of D-H pairs mutually overlap, density carriers (D and H) can move independently of the partners of the D-H pairs released from the binding. In the insulating state ($\xi_{\mathrm{dh}} < \xi_{\mathrm{dd}}$), because most domains of D-H pairs are detached from one another, density fluctuation is localized within the domain of ξ_{dh}. The Mott transition takes place, when the relation $\xi_{\mathrm{dh}} = \xi_{\mathrm{dd}}$ is satisfied.

(6) By adding a D-D (and H-H) repulsive factor, the variational energy is improved slightly, especially in the insulating regime. However, we have not recognized qualitative influences thereof on the Mott transition.

Since the properties of the Mott transition studied in this paper basically coincide with those for the electron systems,[32] the renewed picture of Mott transitions given in §5 can be applied to electron systems.[34] An important remaining problem is to determine the Mott critical value U_c/t more accurately. This problem is closely connected with the great system-size dependence of the present trial wave functions around U_c/t. To remedy it, we may need more exquisite size-dependent correlation factors in the Mott critical regime.[54]

Appendix: Variational Monte Carlo Method

We briefly note the setup and condition of VMC calculations carried out in this paper.

Because we need to optimize variational parameters up to the maximum number L, we use a correlated-measurement or optimization-VMC technique.[19] In the non-linear minimization process of energy expectation values, we adopt a quasi-Newton method, in which gradient vectors are effectively calculated, especially for bosons, by recently proposed formulae,[55] and Hessian matrices are approximately given by Broyden-Flecher-Goldfarb-Shanno formula,[56] the use of which does not affect the accuracy of optimization itself. In coding, we refer to an algorithm offered by Ibaraki and Fukushima.[57] For wave functions with a few parameters, we use a simple linear optimization together.

In both algorithms, parameters as well as energy converges typically after first several rounds of iteration with different fixed sample sets; in each set we generate 2.4-2.5 $\times 10^5$ particle configurations with Metropolis algorithm. After this convergence, we continue excess rounds (15-90 times) of iteration in the optimization process with successively renewed configuration sets. We determine the optimized values by averaging the data obtained in the excess rounds; in averaging, we exclude scattered data beyond the range of twice the standard deviation. Thus, the optimal value is substantially an average of more than several million samples. The variational energy and significant parameters [g and $f(2)$ etc.] are determined with sufficient accuracy in most cases, but accurate determination of insignificant parameters [$f(r)$ with $r > 7$ etc.] is difficult, because E depends on them only very slightly, in other words, particle configurations determining them appear extremely rarely. Anyway, such parameters have little influence on E and other quantities. Physical quantities are calculated with 2.4-2.5 $\times 10^5$ renewed configurations generated by the optimized parameter sets.

Since in Mott critical regimes, the global minimum becomes more competitive with other minima as L increases, accurate energy minimization sometimes becomes not easy, especially for the triangular lattice with QWF and for $\Psi_{\mathrm{DH}}^{\mathrm{opt}}$ and Ψ_{J}. This is the cause of scattered data points near U_c/t in some figures.

1) For instance, M. P. A. Fisher, P. B. Weichman, G. Grinstein and D. S. Fisher: Phys. Rev. B **40** (1989) 546.
2) M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch and I. Bloch: Nature **415** (2002) 39.
3) T. Stöferle, H. Moritz, C. Schori, M. Köhl and T. Esslinger: Phys. Rev. Lett. **92** (2004) 130403.
4) M. Köhl, H. Moritz, T. Stöferle, C. Schori and T. Esslinger:

J. Low Temp. Phys. **138** (2005) 635.

5) I. B. Spielman, W. D. Phillips and J. V. Porto: Phys. Rev. Lett. **98** (2007) 080404.

6) D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner and P. Zoller: Phys. Rev. B **81** (1998) 3108.

7) I. Bloch, J. Dalibard and W. Zwerger: Rev. Mod. Phys. **80** (2008) 885.

8) For one dimension, E. H. Lieb and F. Y. Wu: Phys. Rev. Lett. **20** (1968) 1445.

9) At unit filling, the critical value is estimated at $U_c/t \sim 3.6$ in one dimension (ref. 10), and at $U_c/t \sim 29.3$ (ref. 11) or 31.3 (ref. 12) for the simple cubic lattice.

10) T. D. Kühner and H. Monien: Phys. Rev. B **58** (1998) R14741.

11) B. Capogrosso-Sansone, N. V. Prokof'ev and B. V. Svistunov: Phys. Rev. B **75** (2007) 134302.

12) Y. Kato, Q. Zhou, N. Kawashima and N. Trivedi: Nat. Phys. **4** (2008) 617.

13) W. Krauth and N. Trivedi: Europhys. Lett. **14** (1991) 627.

14) S. Wessel, F. Alet, M. Troyer and G. G. Batrouni: Phys. Rev. A **70** (2004) 053615.

15) B. Capogrosso-Sansone, S. G. Söyler, N. Prokof'ev and B. Svistunov: Phys. Rev. A **77** (2008) 015602.

16) N. Elstner and H. Monien: Phys. Rev. B **59** (1999) 12184.

17) W. L. McMillan: Phys. Rev. **138** (1965) A442.

18) H. Yokoyama and H. Shiba: J. Phys. Soc. Jpn. **56** (1987) 1490.

19) C. J. Umrigar, K. G. Wilson and J. W. Wilkins: Phys. Rev. Lett. **60** (1988) 1719.

20) M. Gutzwiller: Phys. Rev. Lett. **10** (1963) 159.

21) W. Krauth, M. Caffarel, J. -P. Bouchaud: Phys. Rev. B **45** (1992) 3137.

22) D. S. Rokhsar and B. G. Kotliar: Phys. Rev. B **44** (1991) 10328.

23) M. Gutzwiller, Phys. Rev. **137** (1965) A1726.

24) W. F. Brinkman and T. M. Rice: Phys. Rev. B **2** (1970) 4302.

25) M. Capello, F. Becca, M. Fabrizio and S. Sorella: Phys. Rev. Lett. **99** (2007) 056402, and Phys. Rev. B **77** (2008) 144517.

26) C. Castellani, C. Di Castro, D. Feinberg and J. Ranninger: Phys. Rev. Lett. **43** (1979) 1957.

27) T. A. Kaplan, P. Horsch and P. Fulde: Phys. Rev. Lett. **49** (1982) 889.

28) P. Fazekas and K. Penc: Int. J. Mod. Phys. **B1** (1988) 1021; P. Fazekas, Physica Scripta T **29** (1989) 125.

29) H. Yokoyama and H. Shiba: J. Phys. Soc. Jpn. **59** (1990) 3669.

30) H. Yokoyama: Prog. Theor. Phys. **108** (2002) 59.

31) H. Yokoyama, Y. Tanaka, M. Ogata and H. Tsuchiura: J. Phys. Soc. Jpn. **73** (2004) 1119.

32) H. Yokoyama, M. Ogata and Y. Tanaka: J. Phys. Soc. Jpn. **75** (2006) 114706.

33) T. Watanabe, H. Yokoyama, Y. Tanaka and J. Inoue: J. Phys. Soc. Jpn. **75** (2006) 074707.

34) T. Miyagawa and H. Yokoyama: to appear in Physica C (2011), and submitted to J. Phys. Soc. Jpn.

35) H. Yokoyama and M. Ogata: J. Phys. Chem. Solids **69** (2008) 3356.

36) H. Yokoyama, T. Miyagawa and M. Ogata: to appear in Phys-

ica C (2011).

37) For instance, T. P. Meyrath, F. Schreck, J. L. Hanssen, C.-S. Chuu and M. G. Raizen: Phys. Rev. A **71** (2005) 041604.

38) In the previous papers, refs. 32, 33, and 35, we carelessly made a mistake in the expression of operators Q. The correct expression is like in eq. (10) in this paper.

39) The parameter μ' for diagonal neighbors (a part of the sites with $r = 2$) works irrespective of the particle configuration in the NN sites ($r = 1$). In this point, the effect of μ' is not identical with the weight $f(2)$ discussed in the item (2).

40) For instance, A. B. Harris and R. V. Range: Phys. Rev. **157** (1967) 295.

41) F. Gerbier, S.Fölling, A. Widera, O. Mandel and I. Bloch: Phys. Rev. Lett. **96** (2006) 090401.

42) P. Cheinet, S. Trotzky, M. Feld, U. Schnorrberger, M. Moreno-Cardoner, S. Fölling and I. Broch: Phys. Rev. Lett. **101** (2008) 090404.

43) X. Lu and Y. Yu: Phys. Rev. A **74** (2006) 063615.

44) B. Capogrosso-Sansone, E. Kozik, N. Prokof'ev and B. Svistunov: Phys. Rev. A **75** (2007) 013619.

45) W. S. Bakr, A. Peng, M. E. Tai, R. Ma, J. Simon, J. I. Gillen, S. Fölling, L. Pollet, M. Greiner: Science **329** (2010) 547.

46) J. F. Sherson, C. Weitenberg, M. Endres, M. Cheneau and I. Bloch: Nature **467** (2010) 68.

47) E. Kapit and E. Mueller: Phys. Rev. A **82** (2010) 013644.

48) J. Mun, P. Medley, G. K. Campbell, L. G. Marcassa, D. E. Pritchard and W. Ketterle: Phys. Rev. Lett. **99** (2007) 150604.

49) I. B. Spielman, W. D. Phillips and J. V. Porto: Phys. Rev. Lett. **100** (2008) 120402.

50) C. Schroll, F. Marquardt and C. Bruder: Phys. Rev. A **70** (2004) 053609.

51) Precisely speaking, because the symmetry between D and H is broken for $n \neq 1$, we should differentiate the correlation factors between D to H and H to D. However, here we use the identical factor for both, because the difference is negligible even quantitatively for small $|1 - n|$, as we checked in ref. 52 for electronic systems.

52) H. Yokoyama, M. Ogata, Y. Tanaka, K. Kobayashi and H. Tsuchiura: in preparation.

53) In the repulsive factor $g^D \mathcal{P}_J$, the onsite repulsive correlation, mainly controlled by g, is affected also by the long-range part of the repulsive factor $\mathcal{P}_J$, namely the parameter set has redundancy. Consequently, the optimized parameter set are sometimes not uniquely determined, leading to large statistical errors. In this case, the form used in ref. 29 will be better. One should avoid the redundancy in the trial function.

54) D. Tahara and M. Imada: J. Phys. Soc. Jpn. **77** (2008) 093703, and J. Phys. Soc. Jpn. **77** (2008) 114701.

55) C. J. Umrigar and C. Filippi: Phys. Rev. Lett. **94** (2005) 150201; S. Sorella: Phys. Rev. B **71** (2005) 241103.

56) For instance, R. Fletcher: *Practical Methods of Optimization* 2nd ed., (John Wily, Chichester, 1987).

57) T. Ibaraki and M. Fukushima: *FORTRAN77 Optimization Programming*, chap. 6 (Iwanami, Tokyo, 1991), [in Japanese].